Course Specification
— (Bachelor)

Course Title: Data Mining	
Course Code: COMP-454	
Program: Bachelor in Computer Science	
Department: Computer Science	
College: College of Computer Science and Information Technology	
Institution: Jazan University	
Version: V2	
Last Revision Date: 07 January 2023	

Table of Contents

A. General information about the course:

1. Course Identification

1. Credit hours: (03)

2. Course type

A.	☐University	☐College	☒ Department	☐Track	☐Others
B.	☒ Required			☐Elective	

3. Level/year at which this course is offered: (Level 11/ Year 4)

4. Course general Description:

The course aims to introduce the concepts and methods of Data mining. The course is intended to provide the students with strong background knowledge and skills needed for the application of Data Mining. It also includes data warehouses, OLAP, OLTP, various data mining algorithms and methods, association rules, classification, pattern mining, clustering, and outlier analysis. The course provides knowledge of the commonly used data mining techniques and will learn about the application of data mining as well as acquire practical skills in using data mining algorithms. The course also covers advanced topics and recent trends in Data Mining.

5. Pre-requirements for this course (if any):

None

6. Pre-requirements for this course (if any):

No

7. Course Main Objective(s):

- Understand the basic concepts of data mining and data warehousing.
- Develop knowledge of the application of data mining and the social impacts of data mining.
- Demonstrate an understanding of the data pre- processing and data visualization techniques.
- Apply various data mining algorithms for various practical problems using tools.

2. Teaching mode (mark all that apply)

No	Mode of Instruction	Contact Hours	Percentage
1	Traditional classroom	60	100%
2	E-learning		
3	Hybrid • Traditional classroom • E-learning		
4	Distance learning (Self-Learning)		

3. Contact Hours (based on the academic semester)

No	Activity	Contact Hours
1.	**Lectures**	28
2.	**Laboratory/Studio**	28
3.	**Field**	--
4.	**Tutorial**	--
5.	**Others (specify)**	4
Total		60

B. Course Learning Outcomes (CLOs), Teaching Strategies and Assessment Methods

Code	Course Learning Outcomes	Code of CLOs aligned with program	Teaching Strategies	Assessment Methods
1.0	Knowledge and understanding			
1.1	Understand the data mining, data warehousing concepts, KDD process, architecture and tools.	K1	Lectures /Presentations Media Lectures	Mid Term Exam / Assignment -1 Final Theory Exam
1.2	Describe the different types of data sets and their characteristics and measuring the central tendency of the data.	K2	Lectures /Presentations Media Lectures	Mid term exam/ Assignment -1/ Final Theory Exam

Code	Course Learning Outcomes	Code of CLOs aligned with program	Teaching Strategies	Assessment Methods
2.0	Skills			
2.1	Analyze different approaches to data mining with various technologies.	S1	Lectures /Presentations Media Lectures Tutorials	Assignment -1/ Final Theory Exam
2.2	Evaluate different methodologies used in data mining.	S2	Lectures /Presentation Media Lectures Tutorials	Assignment-1/ Final Theory Exam
2.3	Apply various data mining techniques to extract useful information from large amounts of data.	S4	Lectures / Presentations / Lab Demonstration	Group Assignment/ Final Lab Exam Final Theory Exam
2.4	Communicate effectively by applying various DM techniques that require a self-directed piece of practical work that requires the application of data mining techniques.	S5	Lectures /Presentations/ Lab / Demonstration Media Lectures	Group Assignment/ Presentation/ Presentation / Final Lab Exam
3.0	Values, autonomy, and responsibility			
3.1	Integrate data mining applications for data analysis using various tools to solve real-world problems of society.	V1	Lectures /Presentations Lab/Demonstration / Media Lectures	Final Theory Exam / Presentation / Group Assignments
3.2	Demonstrate the ability to work in a team to implement various data mining algorithms and techniques using open-source data mining tools.	V2	Lab Session / Mini project development and demonstration	Presentation / Group Assignments

C. Course Content

No	List of Topics	Contact Hours
1.	Introduction: - What is data mining? Data mining: an essential step in knowledge discovery, diversity of data types for data mining, mining	2T + 2P

	various kinds of knowledge, Data mining: a confluence of multiple disciples, data mining and applications, data mining, and society.	
2.	Data and Measurements- Types of Data sets, Characteristics of Structured data, Attributes, Attributes types, Measuring the Central Tendency, Mean, Median, and mode, Data Compression, Data cube Aggregation, Normalization, Discretization, Data Discretization methods.	2T + 2P
3.	Data Warehousing and Analytical Processing- Sampling- types of sampling. Data warehouse: what and why?, Architecture of data warehouses, OLTP vs OLAP, Three Data Warehouse models, ETL, metadata repository, Data lake, Layers of storage, Conceptual Architecture, data marts, OLAP, and Other operations.	4T + 4P
4.	Pattern mining basic concepts and methods: - What are patterns? What is Pattern Discovery?, K-itemsets and their support. Basic Concepts, Frequent item sets (patterns) and association rules, Efficient pattern mining methods, The Apriori algorithm, Apriori: A candidate generation & test approach, Apriori algorithm (pseudo-code), Direct Hashing and Pruning, Frequent patterns by pattern growth, FPGrowth, Apriori pruning and scalable mining methods, apriori algorithm tricks, candidate generation, Direct Hashing and Pruning	5T + 5P
5.	Classification:- Basic concepts, Bayes Theorem, Naïve Bayes classifier, Lazy Vs Eager Learning, k- Nearest neighbor algorithm.	4T + 4P
6.	Cluster Analysis:- A quick overview, Value of cluster analysis, broad application of cluster analysis, what is cluster analysis, applications Clustering different types of data, User insights and interactions in clustering, Partitioning algorithms, K means clustering method. Hierarchical clustering: basic concepts, Agglomerative clustering, Divisive Clustering,	5T + 5P
7.	Outliers Detection & Data mining research frontiers:- Basic Concepts, Outlier Detection Vs Clustering, What are Outliers?., Outliers Versus Noise, Types of Outliers- Global, Contextual, Collective Outliers Future Trends:- Mining Rich Data types- mining Text Data, mining spatial data, mining graphs and networks, Data mining applications- data mining for sentiment and opinion, data mining for social good, truth discovery and misinformation identification, information and disease propagation, productivity, and team science.	4T+4P
Total		26T+26P

D. Students Assessment Activities

No	Assessment Activities *	Assessment timing (in week no)	Percentage of Total Assessment Score
1.	Midterm Exam (First Two Chapters or ~30% Course)	7th -8th week	15%
2.	Assignment I	9th week	10%
3.	Assignment II (Group assignment)	12th week	15%
4.	Lab Exam	14th Week	20%
5.	Final Theory Exam	12th Week	40%

*Assessment Activities (i.e., Written test, oral test, oral presentation, group project, essay, etc.).

E. Learning Resources and Facilities

1. References and Learning Resources

Essential References	Jiawei Han, Jian Pei, Hanghang Tong, "Data Mining Concepts and Techniques", Fourth Edition, Elsevier, 2023, ISBN: 978-0-12-811760-6.
Supportive References	• David Hand, HeikkiMannila and Padhraic Smyth, "Principles of Data Mining", The MIT Press, 2001, ISBN-10: 026208290X, ISBN-13: 978-0262082907.. • Pang-Ning Tan, Michael Steinbach, Vipin Kumar, "Introduction to Data Mining", Pearson, 2016, ISBN: 9789332571402. • Ian H. Witten, Eibe Frank, Mark A. Hall, Christopher J. Pal, Morgan Kaufmann, "Data Mining: • Practical Machine Learning Tools and Techniques", Fourth edition, 2016, ISBN10: 0128042915, ISBN-13: 978-0128042915.
Electronic Materials	• Blackboard:https://lms.jazanu.edu.sa/webapps/portal/execute/tabs/tabAction?tab_tab_group_id=_1_1 • https://www.kdnuggets.com/data_mining_course/ • https://www.mygreatlearning.com/academy/learn-for-free/courses/data-mining1
Other Learning Materials	

2. Required Facilities and equipment

Items	Resources
facilities (Classrooms, laboratories, exhibition rooms, simulation rooms, etc.)	• Classroom equipped with projector, whiteboard, and sufficient seating arrangements. • Lab with software installed and an individual computer terminal for each student.

Items	Resources
Technology equipment (projector, smart board, software)	• Whiteboards and projectors for classroom and labs • Computer Lab equipped with 30 PCs having Python Anaconda, Google Colab etc • An active internet connection.
Other equipment (depending on the nature of the specialty)	**None**

F. Assessment of Course Quality

Assessment Areas/Issues	Assessor	Assessment Methods
Effectiveness of teaching	Students	Indirect (Course evaluation survey form)
Effectiveness of Students assessment	CRC / QAU / HoD	Direct (Course reports/result analysis)
Quality of learning resources	Track leaders / CRC	Indirect (Review, meetings, and star rating with suggestions for further modification and improvements)
The extent to which CLOs have been achieved	CRC / QAU	Direct (CLO assessment template further verified at course coordinator, Track leader, and QAU level)
Other		

Assessors (Students, Faculty, Program Leaders, Peer Reviewers, Others (specify)
Assessment Methods (Direct, Indirect)

G. Specification Approval

COUNCIL /COMMITTEE	DEPARTMENT COUNCIL
REFERENCE NO.	
DATE	15/10/2022